New edition of Eurachem/CITAC guide on measurement uncertainty arising from sampling

Article (Published Version)

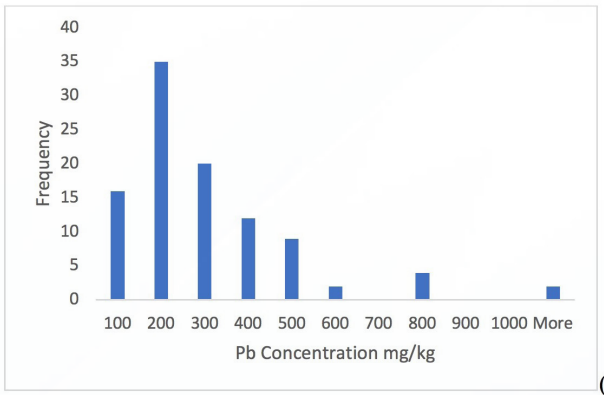
NEW EDITION OF EURACHEM/CITAC GUIDE ON MEASUREMENT UNCERTAINTY ARISING FROM SAMPLING

Michael H. Ramsey // University of Sussex, UK. Chair of Eurachem UfS Working Group

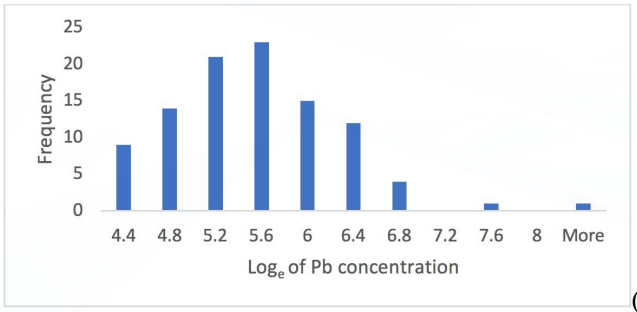
The second edition of the Eurachem/CITAC Guide on measurement uncertainty arising from sampling (UfS) has recently been published [1]. It has been updated to explain how several new research ideas can be used to improve the way that we estimate and express UfS. The Guide considers the whole measurement process, which usually begins at the point where a primary sample is taken. It retains the basic structure of the first edition, describing the estimation of UfS by both empirical and modelling approaches, and includes six worked examples across several application sectors, including food, animal feed, soil and water.

THE UNCERTAINTY FACTOR

One significant new development is the option of using the Uncertainty Factor as an alternative way to express measurement uncertainty. The upper and lower confidence limits of a measurement value are expressed by multiplying and dividing the measurement value by the uncertainty factor, rather than by the traditional approach of adding and subtracting the uncertainty. This approach is more accurate when the relative expanded uncertainty value is large, typically over 20%, and also where the frequency distribution of the uncertainty is approximately log-normal rather than normal. These two conditions often apply to measurement uncertainty that arises from the sampling process, particularly when the spatial distribution of the analyte in the test material is substantially heterogeneous. The Guide explains how the expanded uncertainty factor (^{F}U) can be calculated as $^{F}U = \exp(2s_G)$, where s_G is the standard deviation of the log-transformed measurement values. An updated worked example, for Pb-contaminated soil, is provided to show how ^{F}U can be evaluated in practice using the 'duplicate' method. Duplicated Pb analyses are made on duplicated samples taken at 10 of the 100 sampling targets placed in a grid across a contaminated land site in the usual way. However, the natural logarithms of the Pb measurement values are taken before the analysis of variance (ANOVA) is made. This log-transformation is necessary because the frequency distribution of the Pb measurements on the 100 sampling targets is approximately log-normal (Fig 1a), but much closer to normal after the transformation (Fig 1b). The frequency



(a)



(b)

Fig 1. Histograms of the Pb measurement values for 100 soil targets shown on (a) the original linear scale, showing positive skew (b) after natural logarithms were taken, showing an approximately normal distribution.

distribution of the measurement uncertainty, as judged by the duplicated samples, is also made closer to normal by this transformation [2].

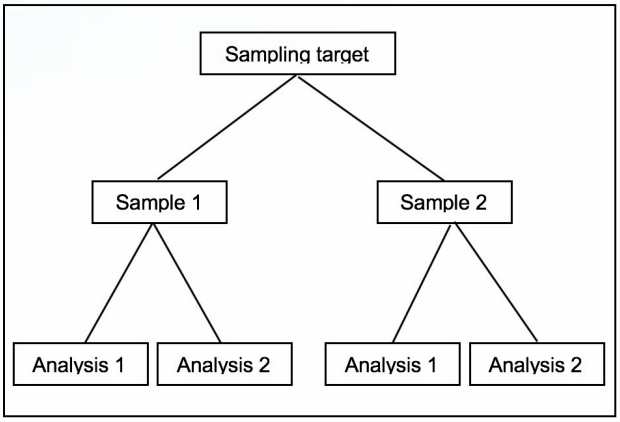
The results of the ANOVA then give not only the expanded uncertainty factor of the measurement ($^{F}U_{meas}$=2.62), but also that arising from the sampling ($^{F}U_{sampling}$=2.60) and from the chemical analysis ($^{F}U_{analysis}$=1.12). The upper confidence limit of a typical Pb measurement value of 300 mg kg^{-1}, can then be calculated as 784 mg kg^{-1} (300 x 2.62), and the lower confidence limit as 115 mg kg^{-1} (300/2.62).

The Guide also explains two options for how measurement uncertainty can be calculated by adding the component arising from sampling, expressed as an uncertainty factor ($^{F}U_{sampling}$), with that arising from chemical analysis, expressed in the traditional way as a relative uncertainty ($U'_{analysis}$). One option is to have both the sampling and analytical uncertainty components calculated and expressed in the log-domain. A second option is to assume, for the analytical component, that the relative standard uncertainty ($s'_{analytical}$) is approximately equal to the standard deviation of the natural logarithms ($s_{G,analytical}$). This is an acceptable approximation when the $s'_{analytical}$ < 0.2, which is usually the case. The two components can then be added as variances in log-space, as in the first option.

AN UNBALANCED EXPERIMENTAL DESIGN TO REDUCE THE COST OF ESTIMATING UfS

A second new development in the methods described in the Guide is the use of an unbalanced experimental design to reduce the cost of estimating UfS by the duplicate method. The first edition of the Guide described the use of a balanced design for the empirical estimation of the measurement uncertainty as a whole, and its two components from the sampling and analytical steps. This balanced design has analytical duplicates on both of the two sample duplicates (Fig 2a). The new edition of the Guide stresses the advantage of using an unbalanced design, with an analytical duplicate on only one of the two sample duplicates (Fig 2b). This design reduces the extra cost of estimating the uncertainty by 33% (i.e. reducing the number of extra measurements required from 3 to 2 per sampling target) [3]. It reduces the number of degrees of freedom on the analytical component, but retains the same number for the sampling component, which is usually the more dominant. Anecdotal evidence suggests that one of the reasons that UfS estimation has not been more widely adopted is the extra cost that is required. This new more economical approach may therefore help to increase the number of application sectors where UfS estimation is fully adopted.

[a]



[b]

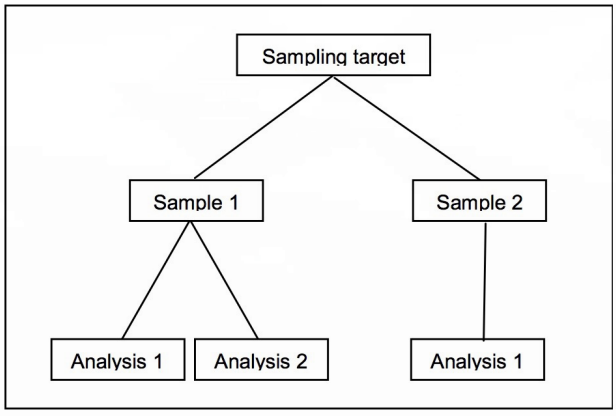


Fig 2: Comparison between (a) the full balanced design, and (b) the unbalanced design proposed in the new edition of the Guide. This reduces the extra cost of estimating the measurement uncertainty due to both the sampling and chemical analysis by 33%.

UfS ESTIMATION USING SAMPLING PROFICIENCY TESTING

A more comprehensive method for the estimation of UfS is by the use of measurements made in Sampling Proficiency Testing (SPT). In the first edition of the UfS Guide this approach was discussed in theory, but the new

edition now refers to the first practical example of the use of SPT data for UfS estimation [4]. In this approach multiple samplers each apply whatever sampling protocol they consider appropriate to achieve the same stated objective for the same sampling target. Using a balanced design across all of the different samplers, it is then possible to included the 'between-sampler' bias in the estimate of UfS, in addition to the components that were previously included. The first practical SPT concerned the measurement of the moisture content of a 20 ton batch of fresh butter. The established duplicate method, using a single sampler, gave an estimate of the expanded relative measurement uncertainty (U') of 0.39%. Interestingly, the use of the SPT results gave an U' estimate of 0.87%, which is a factor of 2.2 larger. This is clear evidence of between-sampler bias, but could conceivably be due to a few poorly performing samplers. When the z-score of the sampler were investigated, there were two samplers that did indeed have potentially non-proficient z-scores ($z > 3$). However, when these two non-proficient samplers were removed from the study, the estimates uncertainty was still 0.69%, which is still 1.8 times larger than the estimate made by the single-sampler approach. The conclusion of this study was that the use of an SPT for the estimation of UfS would give a more reliable, and probably larger estimate of the uncertainty. The multi-sampler approach using an SPT is clearly more expensive to undertake, but in situations where the UfS value has a large economic consequence, such as the estimation of the mass of gold in a potential new mine, it may well be financially justified.

APPLICATION OF UfS ESTIMATION TO NEW SITUATIONS

Another development covered in the new Guide is the increased range of measurement situations where UfS estimation has been applied and reported. One such application is for measurements made *in situ*, where the test material has not been removed from its original location but measured in place. For *in situ* measurements, the taking of a sample is almost indivisible from the rest of the measurement process. This means that situation is more complex than for traditional *ex situ* measurements made in an external laboratory, partially because of the spatial heterogeneity of the analyte concentration in the test material. Even when an *in situ* measurement probe is placed at the same nominal location on the sampling target, the analyte heterogeneity will thereby increase measurement uncertainty. This is due to the higher level of UfS that is present in an *in situ* measurement when compared against an *ex situ* measurement, where the test material has been homogenized. UfS can be estimated for *in situ* measurements using an empirical approach such as the duplicate method [5]. Duplicated positioning of the measurement probe, using the same sampling protocol, can be used to give a 'sample duplicate'. Similarly, duplicated measurements made without moving the probe can be used to give an 'analytical duplicate'. The systematic component of the uncertainty cannot be estimated only with measurements on a matrix-matched certified reference materials (CRMs). This is because a CRM is usually a dried, ground, homogenized and often compacted material that is physically very different from the test material, which may well be moist, unground, heterogeneous and un-consolidated in the measurement situation. Comparison will also be required, therefore, between the measurements made *in situ* and those made *ex situ*, with an independent analytical method for the same measurand, on samples taken from the same sampling target.

'On site' measurements, are made when a sample has been taken from its original location, and usually prepared and homogenised, but it is measured close to its original location. This situation is intermediate in complexity between the traditional *ex situ* measurements, and the *in situ* measurements just discussed. For an example of the determination of total petrol hydrocarbons (TPH) in stockpiled soil, the *ex situ* measurement made by the on-site method can be compared against those made in a remote laboratory under more controlled conditions. Problems can arise from differences between the definitions of the measurand for the two analytical methods for TPH, and in deciding which is correct [6].

UfS IN PASSIVE MEASUREMENTS OF RADIOACTIVE DECAY.

Application of UfS estimation to the measurement of

^{137}Cs in soil, by gamma ray spectrometry at a nuclear decommissioning site, has illustrated some interesting new issues [7]. The passive nature of sampling for *in situ* use of γ-ray spectrometry, means that a very large mass of test material (e.g. 200 - 1000 kg) can form the 'test portion' of this analytical method. This contrasts with the very limited mass of test material (e.g. ~ 0.5 kg) usually physically extracted for the *ex situ* measurement by γ-ray spectrometry. When the duplicate method is applied, the measurement uncertainty from sampling (UfS) for the *in situ* measurements was found to be much lower than that for the *ex situ* measurements. This is undoubtable due to the much greater mass interrogated by the *in situ* γ-ray spectrometry, which is therefore much more representative of the sampling target. This effect is slightly offset by the 50% lower analytical component of the uncertainty, due to the longer counting time typically used for *ex situ* determinations, making the overall expanded measurement uncertainty comparable at around 40%. However, the cost of each *in situ* measurement is about one tenth of an *ex situ* measurement, so it is economically justified to take many more *in situ* than *ex situ* measurements. Overall, when four times more *in situ* are made than *ex situ* measurements, it was found that the standard error on the mean value of ^{137}Cs for the whole site is reduced by a factor of two using the *in situ* measurement approach, at half the cost.

UfS ESTIMATION AT THE MICRO SCALE

The final area of new application of UfS estimation is to a range of different spatial scales. This is particularly the case for instrumental measurements made using 'beam sampling' at scales ranging from the millimetre to the micron scale [8]. At these smaller scales, analyte heterogeneity become increasing important. The heterogeneity is often the main component of the UfS and hence the dominant source of the measurement uncertainty. Studies using PXRF at the millimetre scale, and SIMS at the micron scale, have used the duplicate method to estimate both the UfS, and the analyte heterogeneity. When the UfS is included in the uncertainty estimate, it is possible to show that these *in situ* measurements can be fit-for-purpose (FFP, such as the spatial mapping of element concentration), despite having higher uncertainty than is usual for bulk analysis. With an increasing use of *in situ* measurement devices in many sectors of society, at all spatial scales, reliable methods of estimating the UfS of beam measurement procedures can enable their FFP to be judged.

REFERENCES

[1] M H Ramsey, S L R Ellison and P Rostron (eds.) Eurachem/EUROLAB/CITAC/Nordtest/AMC Guide: *Measurement uncertainty arising from sampling: a guide to methods and approaches.* Second Edition, Eurachem (2019). ISBN (978-0-948926-35-8). Available from http://www.eurachem.org

[2] Analytical Methods Committee (2019). Why do we need the uncertainty factor? Technical Brief 88. DOI: http://dx.doi.org/10.1039/c9ay90050k Anal. Methods, 2019, 11, 2105–2107, https://pubs.rsc.org/en/content/articlelanding/2019/ay/c9ay90050k#!divAbstract

[3] P Rostron and M H Ramsey (2012) Cost effective, robust estimation of measurement uncertainty from sampling using unbalanced ANOVA. Accreditation and Quality Assurance: Journal for Quality, Comparability and Reliability in Chemical Measurement. 17, 1, 7-14. DOI: 10.1007/s00769-011-0846-2, http://www.springerlink.com/content/c54k251654534876/fulltext.pdf

[4] M H Ramsey. B Geelhoed, A P Damant, R Wood (2011) Improved evaluation of measurement uncertainty from sampling by inclusion of between-sampler bias using sampling proficiency testing. Analyst, 136 (7), 1313 – 1321. DOI:10.1039/C0AN00705F

[5] M H Ramsey and K A Boon (2012) Can *in situ geochemical measurements be more fit-for-purpose than those made ex situ? Applied Geochemistry 27, 969-976.* http://dx.doi.org/10.1016/j.apgeochem.2011.05.022

[6] K A Boon and M H Ramsey (2012) Judging the fitness of on-site measurements by their uncertainty, including the contribution from sampling. *Science of the Total Environment 419, 196–207, http://dx.doi.org/10.1016/j.scitotenv.2011.12.001*

[7] *P Rostron, J A Heathcote, M H Ramsey (2015) Evaluation of uncertainties in in situ and ex situ gamma measurements on land areas with low contamination levels. Journal of Radiological Protection 35 (2015) 391-399. DOI:* https://dx.doi.org/10.1088/0952-4746/35/2/391

[8] Analytical Methods Committee (2018) AMC Technical Brief No 84. Beam sampling: taking samples at the micro-scale, Analytical Methods, 10, 1100-1102